

EUROPEAN PATENT OFFICE

Patent Abstracts of Japan

PUBLICATION NUMBER : 04062600
PUBLICATION DATE : 27-02-92

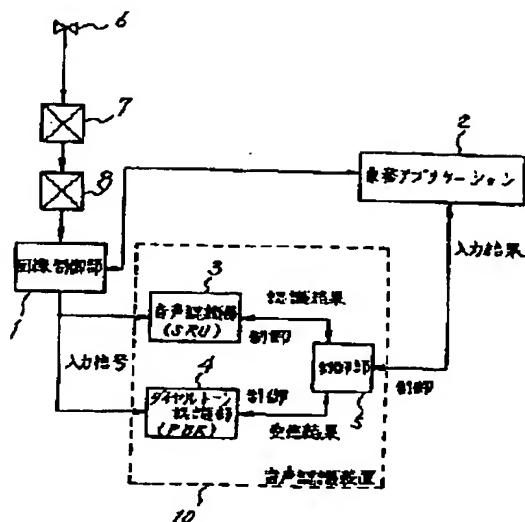
APPLICATION DATE : 02-07-90
APPLICATION NUMBER : 02174801

APPLICANT : NEC CORP;

INVENTOR : HAYASHI HARUYUKI;

INT.CL. : G10L 5/06 H04M 11/10 H04Q 1/45

TITLE : VOICE RECOGNITION DEVICE



ABSTRACT : PURPOSE: To eliminate need to divide telephone lines of a system side into lines for PB input and lines for voice input by providing a voice recognition part which discriminate a voice from an input signal, a PB reception part which detects a PB signal from the input signal, and a control part which discriminates between the voice signal and PB signal automatically.

CONSTITUTION: When a user makes a telephone call through a telephone set 6, the call is sent to a line control part 1 through exchanges 7 and 8. The control part 1 informs service application 2 that there is the call from the user and connects the line to a voice recognition device 10 once receiving a termination instruction from the application 2. The application 2 instructs the number of input signal words to be recognized to a control part 5 which controls the device 10 in response to the termination completion report from the control part 1. The control part 5 receives a digit quantity instruction from the application 2, actuates a voice recognition part 3 and a PB signal detection part 4 on receiving the digit quantity instruction from the application 2, and performs simultaneous recognition as to an input signal supplied from the control part 1, thereby deciding the input signal automatically according to the result. After a specific number of digits are recognized, the recognition result is outputted to the application 2.

COPYRIGHT: (C)1992,JPO&Japio

⑫ 公開特許公報(A) 平4-62600

⑮ Int. Cl.⁵

識別記号

庁内整理番号

⑬ 公開 平成4年(1992)2月27日

G 10 L 5/06
H 04 M 11/10
H 04 Q 1/45B 8842-5H
7117-5K
Z 8843-5K

審査請求 未請求 請求項の数 1 (全7頁)

⑭ 発明の名称 音声認識装置

⑯ 特 願 平2-174801

⑰ 出 願 平2(1990)7月2日

⑱ 発 明 者 林 晴 之 東京都港区芝5丁目7番1号 日本電気株式会社内

⑲ 出 願 人 日本電気株式会社 東京都港区芝5丁目7番1号

⑳ 代 理 人 弁理士 内 原 晋

明 細 書

1. 発明の名称

音声認識装置

2. 特許請求の範囲

- (1) 入力信号から音声認識して認識結果を出力する音声認識部と、前記入力信号からPB信号を検出して検出結果を出力するPB信号検出部と、前記音声認識部および前記PB信号検出部を制御する制御部とを有する音声認識装置において、

前記制御部が、前記音声認識部と前記PB信号検出部を同時に用いたときの前記入力信号の認識結果および検出結果に基づき、前記入力信号が音声かPB信号かを自動判別することを特徴とする音声認識装置。

- (2) 前記制御部が、

前記音声認識部の応答である第1の応答が前記PB信号検出部の応答である第2の応答より

も先で、かつ前記第1の応答から所定時間内に前記第2の応答がない場合、前記入力信号を音声と判別し、前記所定時間内に前記第2の応答があった場合および前記第2の応答が前記第1の応答よりも先の場合に前記入力信号をPB信号と判別することを特徴とする特許請求の範囲第1項記載の音声認識装置。

- (3) 前記制御部が、

前記音声認識部の応答である第1の応答よりも前記PB信号検出部の応答である第2の応答が先で、かつ前記第2の応答から第1の所定時間内に新たな前記第2の応答があった場合および、前記第2の応答よりも前記第1の応答が先で、かつ前記第1の応答から第2の所定時間内に前記第2の応答があり、この第2の応答から前記第1の所定時間内に新たな前記第2の応答があった場合および

前記入力信号の認識を複数回行なった結果、前記第2の応答が所定回数以上あった場合に前記入力信号をPB信号と判別することを特徴と

する特許請求の範囲第1項記載の音声認識装置。

3. 発明の詳細な説明

〔産業上の利用分野〕

本発明は、電話回線を介した音声認識応答システム等に用いられている音声認識装置に関し、特にPB信号検出機能(PB受信機能)を有する音声認識装置に関する。

〔従来の技術〕

従来、この種の音声認識装置は音声認識部(SRU)とPB信号検出部(PBR)とを有しているが、電話回線からの入力信号がPB信号(PB)か音声かを自動的に区別できなかった。そのためPB入力用と音声入力用の電話回線を別個に設け、音声認識装置を利用するソフトウェア等の業務アプリケーションにおいて、どちらの回線に電話がかかったかを監視し、対応する認識部のみを用いるよう音声認識装置に指示を行っていた。

〔発明が解決しようとする課題〕

上述した従来のPB受信機能を有する音声認識

装置(以下、単に音声認識装置)を用いる場合、音声入力用回線とPB入力用回線を独立して設ける必要がある。また、業務アプリケーションにおいて回線に応じた処理を行わなければならないほか、ユーザにおいては音声入力、PB入力により別の電話番号を用いることになり、サービス上の問題もあった。さらに、発信を装置側から行う必要のある場合、ユーザが音声入力とPB入力のいずれを用いるのか、予め登録する必要があった。また、回線と入力信号の種別対応が固定なため、空いた回線の効率的割り当てができず、音声入力用回線又はPB入力用回線の一方に呼が集中した場合、ユーザに対するサービスが低下するという問題もあった。

〔課題を解決するための手段〕

本発明の目的は、上述した課題を解決した音声認識装置を提供することにある。このため、本発明においてはSRUとPBRの両方で入力信号の認識を行ない、各認識部の応答結果をもとに入力信号を音声かPBか自動判別することにより上記

目的を達成している。

〔実施例〕

次に、本発明について図面を参照して説明する。

第1図は本発明の音声認識装置を利用した音声認識システムの一実施例である。第1図において、ユーザの利用する電話器6は、第1および第2の交換機7および8を介して回線制御部1に接続されている。回線制御部1は業務アプリケーション2の制御に従って入力信号を音声認識装置10に与える。音声認識装置10は音声認識部(SRU)3、PB信号検出部(PBR)4および制御部5から構成され、業務アプリケーション2に制御される。次に、第2図に示すフローチャートを参照して動作について説明する。第1図において、本システムを利用するユーザが電話器6により発呼すると、第1および第2の交換機7および8を介して回線制御部1に発呼が伝わる(ステップ201)。回線制御部1はユーザからの電話がかかってきたことを業務アプリケーション2に知らせ、業務アプリケーション2の着信指示を受ける

と、回線を音声認識装置10に接続する(ステップ202)。業務アプリケーション2は回線制御部1の着信完了通知に対応して、音声認識装置10を制御する制御部5に認識すべき入力信号語数を指示する(ステップ203)。本実施例で同時認識すべき入力信号は電話器のPBダイヤルの種類に等しく、そのほとんどは数字であるため、語数を桁数として以降取り扱う。

音声認識装置10の制御部5は、業務アプリケーションからの桁数指示を受けてSRU3およびPBR4を起動し(ステップ204)、回線制御部1から供給される入力信号について同時認識を行い、その結果により、入力信号の自動判別を行う(ステップ205)。所定の桁数分の認識が終了したら、業務アプリケーション2に認識結果を出力する(ステップ206)。

次に本発明の特徴である入力信号の自動判別を更に詳細に説明する。入力信号の自動判別は、以下の条件に基づいている。

(1) PBRの検出率はほぼ100%であるがSRU

の認識率はこれより低い。

(2) 音声とPBを混在利用するユーザはいない。

第3図は、第2図における自動判別ステップ205の詳細なフローチャートで音声認識装置10の制御部5がこのフローチャートに従い動作する。制御部5はたとえばマイクロプロセッサにより実現される。入力信号の1桁目については同時認識を行うため、SRU3およびPBR4を起動する(第2図ステップ204)。次に制御部5は内部のタイマに所定時間(T1)を設定する。

もし、先にPBR4から1桁目の受信結果が返れば(ステップ302)、PBR4の認識率は100%であると考えているため、即これを1桁目の同時認識の結果とし(ステップ311)、SRU3をこの時点でストップさせる。複数桁入力の場合上記(2)の条件にもとづき、2桁目以降に音声で入力することはないと考え、2桁目以降は同時認識でなくPBR4のみを用いて処理を行う(ステップ312)。

もし、1桁目の結果が先にSRU3から返れば

クする(ステップ304)。T1時間経過していれば業務アプリケーション2にタイムアウトした旨を知らせて処理を終了する(ステップ306)。T1時間経過していなければ再びステップ302へ戻り、PBR4およびSRU3の応答有無をチェックする。以上の動作により、入力信号の自動判別が実現される。

この第1実施例では1桁目の認識結果に基づいて2桁目からの認識をSRU又はPBR一方で行っていた。この方法は入力信号に対応する認識部のみが正常に回答するほとんどの場合には問題なく動作する。しかしながら実際にはごくまれに、両方の認識部が回答することがあり、同時認識により生じる問題である。この場合、第1の実施例では誤認識の可能性がある。両方の認識部が回答するのは、

- (a) 音声入力に対してPBRも応答する、
 - (b) PB入力に対してSRUも応答する、
- 2通りがある。(a)は頻度は小さいが不可避である。このため(a)の事象が連続して発生する可能性は低

(ステップ303)、本当にPBR4からの結果が返らないかをある時間(T2)だけ待つ。

このため、制御部5はT1を設定したタイマとは別のタイマにT2を設定する(ステップ307)。

この待ち時間T2は処理時間に遅れを与える為、長くとも1.5秒程度にしておく。この間にPBR4応答があれば(ステップ308)、SRU3は雑音等で誤認識したと見なし、PBR4からの結果を1桁目の同時認識の結果とし、SRU3はこの時点でストップさせる(ステップ311)。複数桁入力の場合、2桁目からはPBR4のみを用いて認識する(ステップ312)。またステップ308でT2時間内にPBR4からの応答がなければ入力信号は音声であると判断し(ステップ309)、2桁目以降の入力信号対してはSRU3のみを用いて認識する(ステップ310)。所定の桁数分の認識処理が終了したら、その結果を業務アプリケーション2へ出力する(ステップ206)。一方、ステップ303においてSRU3の応答がなければT1時間経過したかをチェッ

いと考えて、予め定めた複数桁でPBRが応答すれば、入力信号はPBであると判断する。(b)については前述した条件(1)からSRUの認識率は100%でないがPBRの認識率は100%と考えられる為、PBと判定する。しかし、PBが入力される前に雑音や音声が入力しこれをSRUが認識した場合、PBRはSRUより遅れて結果を返す事がある。この解決策として、SRUからの結果が返ってもPBRからの結果が返るのをしばらく待つという方法があるが、この待ち時間は認識応答時間を遅らせることになる為あまり長い時間を設定することができず不十分である。そこで、もう一つの解決策として、1桁ずつSRUからの結果が返るたびにPBRの結果をみて、2桁以上PBRから結果が返れば間違いなくPBであると見なすことでより正確な自動判別が可能となる。第4図は、この原理に基づく第2の実施例を示すフローチャートである。

第4図は、第1実施例を示す第3図と同様に、第2図におけるステップ205に対応する。すな

わち、SRUとPBRはいずれも起動されている。まず、制御部5は、業務アプリケーション2の指示した認識桁数に応じた入力時間T0を第1のタイマに設定する。

もし、先にPBRから1桁目の受信結果が返れば(ステップ402)、PBRの認識率は100%であると考え、即これを1桁目の同時認識の結果とする(ステップ411)。もし、1桁目の結果が先にSRUから返れば(ステップ403)、本当にPBRからの結果が返らないかをある時間(T1)だけ待つ(ステップ406)。この待ち時間は入力処理時間に遅れを与える為、長くても1.5秒程度にしておく。この間にPBRからの結果の結果が返れば、SRUは雑音等で誤認識したと考え、PBRからの結果を1桁目の同時認識の結果とする(ステップ411)。

1桁目がPBだと判断した場合(ステップ411)、PBRの応答回数は無条件に1回だからステップ412は「NO」の方向に進み、次桁の処理に移る(ステップ413)。この場合、SRUの起動は

る(ステップ407)。

本実施例で複数桁の認識処理を行う場合、PBRとSRUの応答が混在するときには、その状況に応じて認識結果のおきかえを行なう。おきかえには、

- (7) 複数桁の認識処理中、2桁以上PBRの応答があった場合、その時点で認識済の結果をPBRの応答結果におきかえる第1のおきかえ、
- (4) ある桁についてPBRの応答結果を認識結果とした後、次桁に対するPBRの応答がT2時間内になく、かつある桁までの認識結果が全てSRUの応答結果である場合、ある桁までの認識結果をSRUの出力結果におきかえる第2のおきかえ

の2通りがある。これら第1および第2のおきかえの発生する具体的な認識処理について、第4図と第5図を併せて説明する。

第5図は、5桁の認識処理実行時におけるSRUとPBRの応答と、認識結果の状態の示す図である。1桁目に対する同時認識の結果、SRUが先

せずPBRからの結果をある時間(T2)だけ待つ(ステップ414)。この待ち時間はユーザがゆっくり電話器のボタンを押す間隔の時間で、たとえば1秒〜3秒程度にしておく。この間にPBRからの結果が返れば、PBRからの結果を2桁目の同時認識の結果とし(ステップ417)、最終桁でなければ(ステップ418)次の桁以降は同時認識でなくPBRのみを用いた認識処理を行う(ステップ419)。逆に、この待ち時間(T2)の間にPBRからの結果が返らず(ステップ414)、且つSRUから1桁目の結果が返っていれば(ステップ415)、SRUからの1桁目の結果を同時認識の1桁目の結果に置き換え(ステップ416)、2桁目としてSRUを起動する(ステップ410)。

一方、1桁目の結果が先にSRUから返り(ステップ403)、待ち時間(T1)の間にPBRからの結果の結果が返らなければ(ステップ406)、PBRは間違わないからPBではないと判断し、SRUからの結果を1桁目の同時認識の結果とす

に回答し(ステップ403)、T1時間内にPBRの応答がない(ステップ406)ため、1桁目の認識結果はSRUの応答結果S1とする(ステップ407および第5図(a))。2桁目も同時認識を行った結果、再びSRUが先に回答し、T1時間内にPBRの応答がないため、2桁目の認識結果もSRUの出力S2とする(第5図(b))。ところが3桁目の同時認識ではPBRがSRUより先に回答し(ステップ402)、3桁目の認識結果はPBRの出力P1となる(ステップ411および第5図(c))。PBRの応答回数は1回なのでステップ412は下へ進み、次桁の認識処理を行う(ステップ413)。ここでT2時間内にPBRから応答がなかったとする(ステップ414)。この時点で制御部5はSRUが3桁目に回答しているかを調べ(ステップ415)、応答があれば3桁目までの認識結果をSRUの出力に全ておきかえる上記(4)のおきかえが発生する(ステップ416、第5図(d))。SRUは1桁の認識ごと

に起動する必要があるので、次桁処理のためにSR

Uを起動する(ステップ410)。次桁に対しては再びPBRが先に応答したとする(ステップ402)。PBRが応答したため、その桁の認識結果はPBRの応答結果とする(ステップ411)。この時点で、認識処理開始時からPBRが2度応答しているため、入力信号はPBと判断し、上記(7)のおきかえが発生する(ステップ412、417および第5図(e))。入力信号をPBと判断したため、次桁以降はSRUを起動せずPBRのみを用いて認識を行ない(ステップ419および第5図(f))、5桁分の認識結果が得られたら(ステップ418および第5図(g))業務アプリケーションに出力する(ステップ206)。本実施例においては、より正確な入力信号の判別が可能となる効果がある。また、第4図ステップ412における条件のPBRの2桁以上に対する応答としたが、複数桁に対する応答とすれば本実施例の効果は得られる。

[発明の効果]

以上説明したように本発明による音声認識装置

は、入力信号から音声を識別する音声認識部と、入力信号からPB信号を検出するPB受信部と、音声信号かPB信号かを自動的に判定することができる制御部を有することにより、本装置を制御する業務アプリケーションがPB信号か音声信号かを意識する必要がなくなるという効果を有する。

したがって、システム側の電話回線をPB入力用と音声入力用にわけの必要がなくなり、結果として、利用客にとっては、例えば照会サービスの場合は音声入力用の電話番号とPB入力用の電話番号を区別する必要がなくなり、通知サービスの場合はシステム側にPB入力か音声入力かを予め登録しなくてもよくなり、サービス性が向上されるという効果がある。

また、システム側にとっては、同一電話回線でPB入力処理と音声入力処理を行うのでトラフィックが効率的になり、システムの処理能力が十分に発揮できるという効果がある。

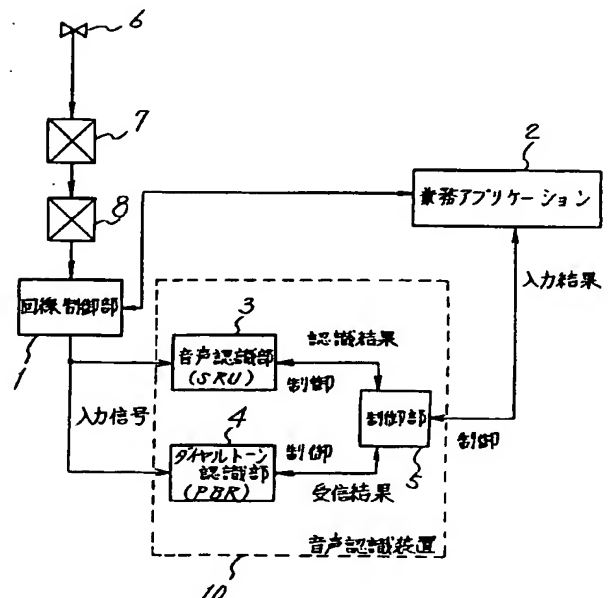
4. 図面の簡単な説明

第1図は本発明による音声認識装置を用いた音声認識システムの一例を示すブロック図、第2図は第1図に示したシステムの動作を説明するフローチャート、第3図～第5図は本発明の音声認識装置における自動判別動作を説明するフローチャートをそれぞれ示す。

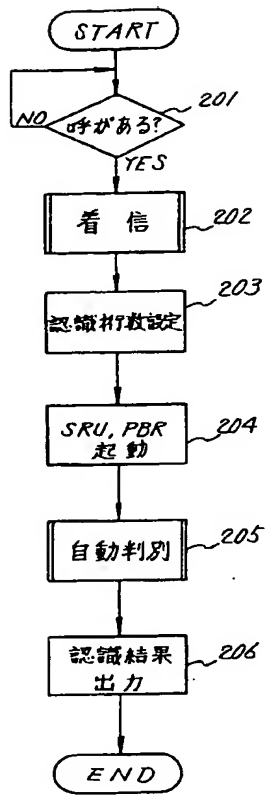
第1図において、

1……回線制御部、2……業務アプリケーション、3……音声認識部(SRU)、4……ダイヤル・トーン認識部(PBR)、5……制御部、6……加入者電話、7、8……交換機、10……音声認識装置。

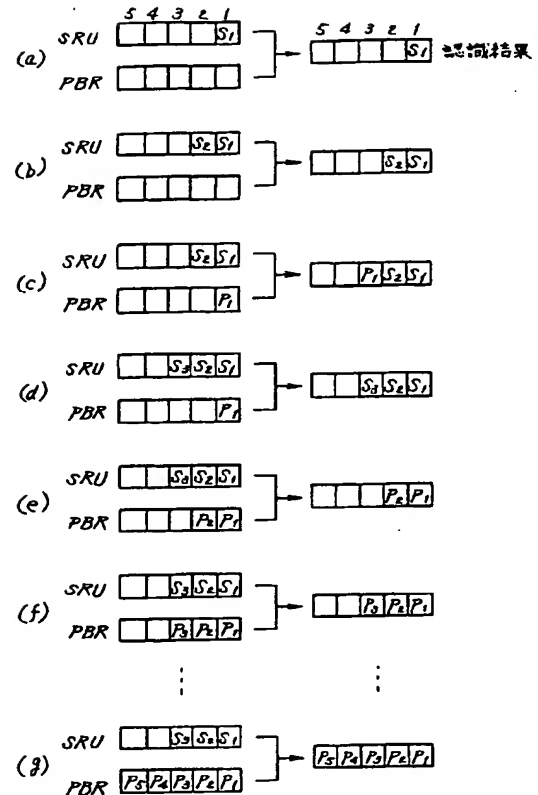
代理人 井理士 内 原 晋



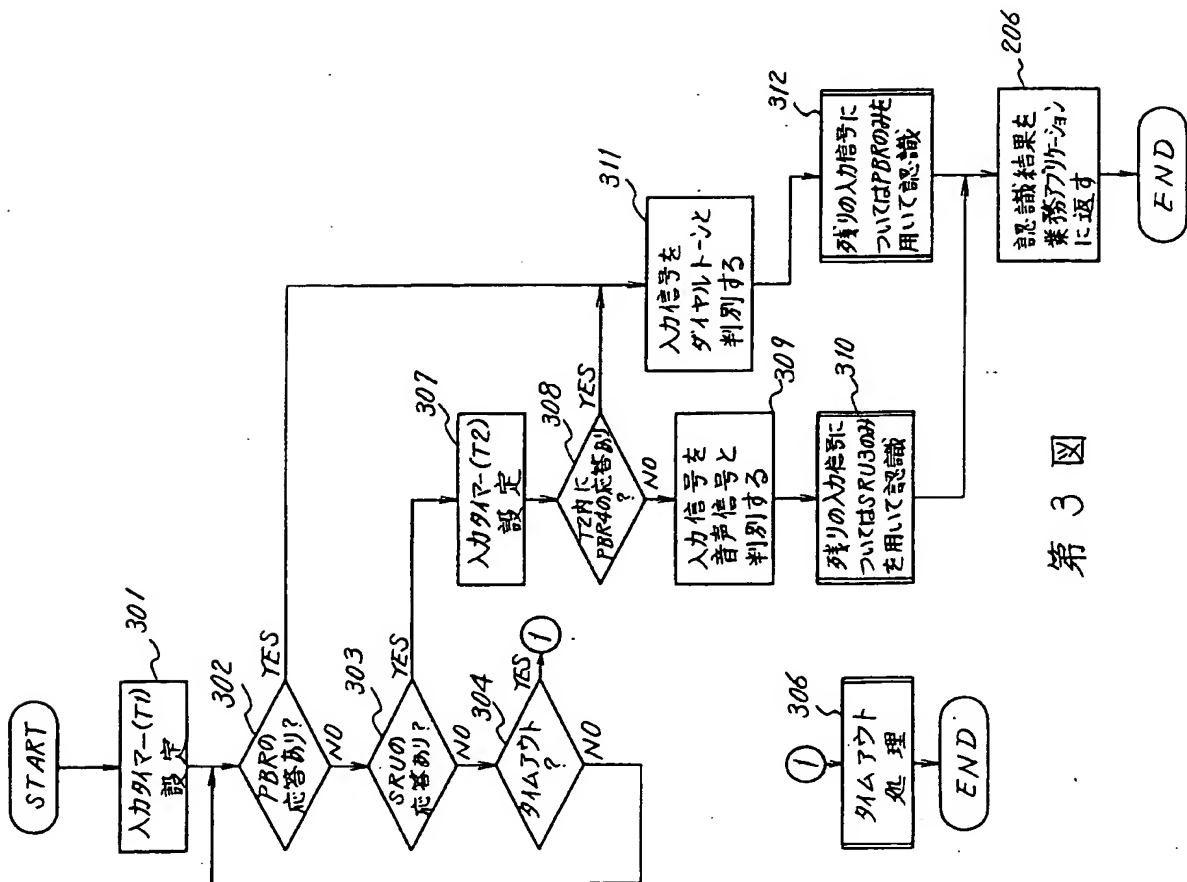
第1図



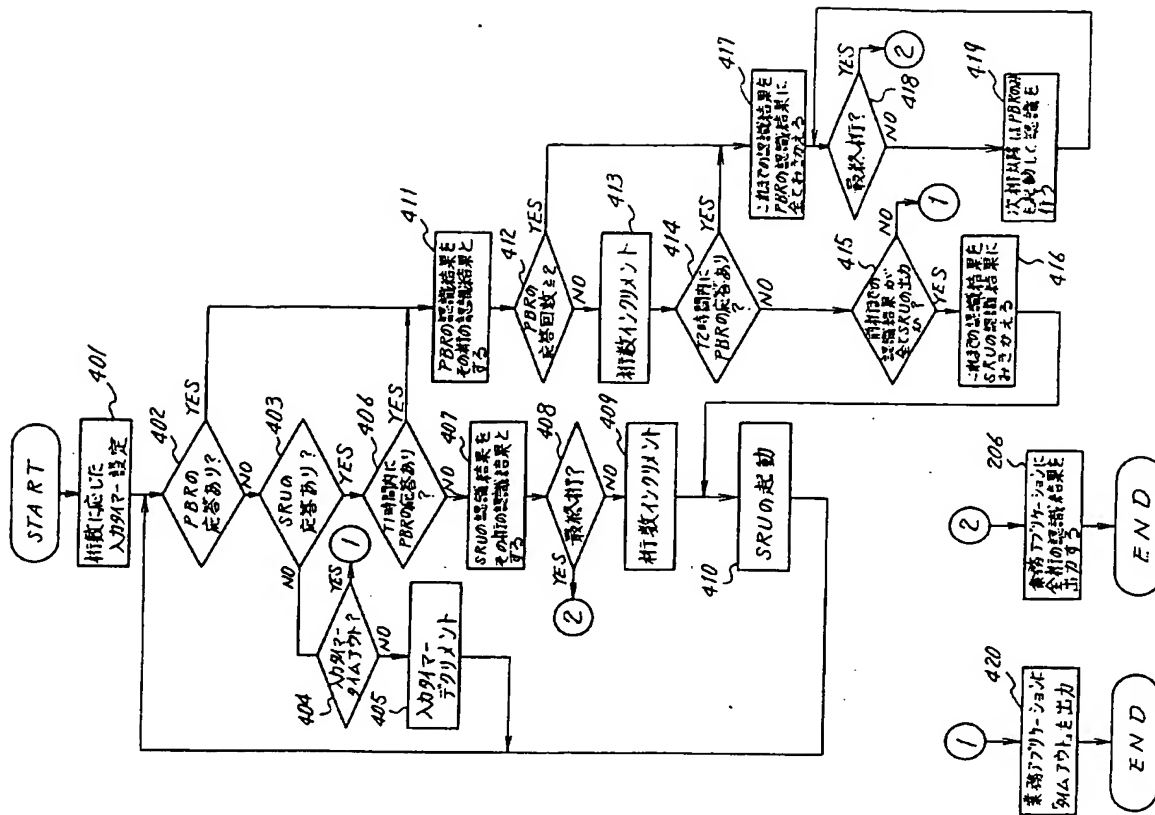
第 2 図



第 5 図



第 3 図



第 4 図